Engineering Cellular Metabolism

Nielsen, Jens; Keasling, Jay

Published in:
Cell

Link to article, DOI:
10.1016/j.cell.2016.02.004

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Leading Edge
Review

Cell

Engineering Cellular Metabolism

Jens Nielsen[1,2,3,*] and Jay D. Keasling[2,4,5,6,7,*]
[1]Department of Biology and Biological Engineering, Chalmers University of Technology, Kemivägen 10, SE412 96 Gothenburg, Sweden
[2]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kogle Allé, DK2970-Hørsholm, Denmark
[3]Science for Life Laboratory, Royal Institute of Technology, SE17121-Solna, Sweden
[4]Joint Bioenergy Institute, Emeryville, CA 94608, USA
[5]Department of Chemical and Biomolecular Engineering & Department of Bioengineering, University of California, Berkeley, Berkeley, CA 94720, USA
[6]Biological Systems & Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[7]Synthetic Biology Engineering Research Center (Synberc), Berkeley, CA 94720, USA
*Correspondence: nielsenj@chalmers.se (J.N.), keasling@berkeley.edu (J.D.K.)
http://dx.doi.org/10.1016/j.cell.2016.02.004

Metabolic engineering is the science of rewiring the metabolism of cells to enhance production of native metabolites or to endow cells with the ability to produce new products. The potential applications of such efforts are wide ranging, including the generation of fuels, chemicals, foods, feeds, and pharmaceuticals. However, making cells into efficient factories is challenging because cells have evolved robust metabolic networks with hard-wired, tightly regulated lines of communication between molecular pathways that resist efforts to divert resources. Here, we will review the current status and challenges of metabolic engineering and will discuss how new technologies can enable metabolic engineering to be scaled up to the industrial level, either by cutting off the lines of control for endogenous metabolism or by infiltrating the system with disruptive, heterologous pathways that overcome cellular regulation.

Introduction

For at least 8,000 years, humans have harnessed microbes to produce fermented foods and beverages. In more recent history, microbes have been used to produce chemicals for a wide range of applications. During World War I, Chaim Weismann developed the acetone-butanol-ethanol fermentation process, which was used for ~50 years to produce acetone and is now being revived for production of 1-butanol. In the 1920s, fermentation of the filamentous fungus *Aspergillus niger* was adapted to generate citric acid, a food and beverage ingredient. During World War II, the same technology was used for industrial scale production of penicillin, the first pharmaceutical produced by fermentation.

The following decades witnessed a dramatic increase in the use of microorganisms to synthesize natural products of pharmaceutical interest, such as antibiotics, cholesterol lowering agents, immunosuppressants, and anti-cancer drugs. Improved performance of classical fermentation processes for such purposes was typically achieved through mutagenesis and screening. For antibiotics in particular, this was an extremely efficient approach, with penicillin production using *Penicillium chrysogenum* boosted by more than 10,000-fold (Thykaer and Nielsen, 2003). Although genetic engineering made it possible to use a more directed approach to improve metabolism, most work focused on the development of cell factories for production of recombinant proteins for use as pharmaceuticals, and today, there are more than 300 biopharmaceutical proteins and antibodies on the market with sales exceeding $100 billion (Langer, 2012).

With the late 1980s and early 1990s came new insights into the complex inner workings of cellular metabolism, fueled by bioinformatics and mathematical modeling methods that allowed quantitative analysis. This enabled specific genetic modifications altering cellular metabolism to be introduced, such that fluxes could be directed toward the product of interest. Thus, the field of metabolic engineering was born (Bailey, 1991; Stephanopoulos and Vallino, 1991; Nielsen, 2001; Keasling, 2010). Now, more than twenty years later, metabolic engineering has been exploited not only to improve traditional microbial fermentation processes, but also to produce chemicals that are currently used as fuels, materials, and pharmaceutical ingredients (Table 1).

Despite the advanced systems and synthetic biology technologies now available for detailed phenotypic characterization of cells and genome editing, developing new cell factories that meet the economic requirements for industrial scale production is still challenging, typically requiring 6–8 years and over $50 million. The reason for this is inherent to the cells themselves. To ensure metabolic homeostasis even when exposed to varying environmental conditions, cells have evolved extensive regulation and complex interactions between metabolic pathways. Redirecting carbon fluxes toward desired metabolites therefore requires modulating the lines of communication in endogenous metabolic pathways or infiltrating the system with disruptive signals that interfere with these regulatory mechanisms. At present, our knowledge of how metabolism is regulated even in simple model cells is limited. As a result, engineering a cell factory involves several rounds of the so-called "design-build-test" cycle, in which a certain metabolic design is implemented and improved through genetic engineering and thereafter tested.

Table 1. Some Success Stories of Metabolic Engineering

Chemical	Application	Cell Factory	Companies
Lysine	feed additive (>1 million tons/year)	*Corynebacterium glutamicum*	Evonik, ADM, CJ, Ajinomoto
1,3-Propanediol	chemical building block, e.g., for production of materials, cosmetics, and food ingredients	*Escherichia coli*	Dupont and Tate&Lyle joint venture
7-ADCA	precursor for the broad-spectrum antibiotic Cephalexin	*Penicillium chrysogenum*	DSM
1,4-Butanediol	chemical building block, e.g., for production of Spandex	*Escherichia coli*	Genomatica
Artemisinic acid	anti-malarial drug	*Saccharomyces cerevisiae*	Sanofi Aventis (process developed by Amyris)
Isobutanol	advanced biofuel	*Saccharomyces cerevisiae*	Gevo, Butamax

Here, we will discuss the principles and current challenges of metabolic engineering, focusing on how metabolism can be engineered for industrial level production of specific chemicals, either through de-regulation of endogenous metabolism or through insertion of heterologous pathways that overcome cellular regulation. We will then discuss how technologies developed in recent years can contribute to the design-build-test cycle and how adding a fourth element to this cycle, namely "learn," can improve the process. Based on implementation of specific metabolic designs, can we gain new knowledge about how metabolism operates and how it is regulated and subsequently use this knowledge for improved design?

Challenges for Metabolic Engineering

Even though metabolic engineering has found applications in optimization of existing processes, much of the current focus is on the development of novel bioprocesses. In the fuel and chemical industry, there is much interest in exploiting the potential of bio-based production for two major reasons: the sustainability factor and the possibility of producing new molecules. Bio-based production of chemicals allows for use of renewable raw materials, such as plant-derived feedstocks like starch, sucrose, cellulose, and lignocellulose that are more sustainable than many traditional chemical processes relying on fossil fuels. Furthermore, replacement of traditional chemical synthesis with bio-based production typically results in reduced environmental footprint in terms of energy usage and emission (Saling, 2005). The key driver for the chemical industry is, however, the production of chemicals that have either better properties than traditional chemicals or chemicals that can find new applications.

The Route for Development of a Novel Bioprocess

Production of a so-called "drop-in" chemical starts with identification of the molecule of interest, followed by determination of whether there exists a metabolic pathway in nature to produce this molecule (Figure 1A). Drop-in chemicals are molecules produced by fermentation instead of from fossil feedstock or other natural sources that are difficult to work with (such as rare plants). In some cases, it is possible to identify a natural producer of the molecule, and this cell factory can then be used for further improvement. If, on the other hand, you want to transfer the biosynthetic pathway to a heterologous host, and if all of the enzymes of the biosynthetic pathway have not yet been identified, heterologous expression requires enzyme discovery as part of the metabolic engineering program, as illustrated for production of artemisinic acid (Ro et al., 2006; Westfall et al., 2012; Paddon et al., 2013) and opioids (Galanie et al., 2015). In some cases, however, it is difficult to identify all the biosynthetic enzymes needed to produce a molecule, and this hinders pathway reconstruction in a heterologous host. For instance, not all the enzymes involved in biosynthesis of the anti-cancer drug taxol have yet been identified (Ajikumar et al., 2010). Improved technologies for DNA and RNA sequencing, bioinformatics, and structure-function predictions have advanced our ability to rapidly identify enzyme candidates for a specific biosynthetic pathway that can subsequently be evaluated for their ability to reconstruct a complete pathway. In case it is not possible to identify a natural producer, chimeric pathways may have to be reconstructed and some of the enzymes may have to be evolved or engineered to have new features.

Traditionally, natural producers were developed for production of the molecule of interest through classical strain improvement. However, with the advent of metabolic engineering, the preferred route for developing a novel bio-process is now through the use of "platform cell factories" (Figure 1A). Examples include *Saccharomyces cerevisiae*, *Escherichia coli*, *Aspergillus niger*, *Bacillus subtilis*, *Corynebacterium glutamicum*, and Chinese hamster ovary (CHO) cells. The advantage of using platform cell factories are numerous: (1) they are very well characterized in terms of genetics and physiology; (2) it is easier to obtain product approval by governmental organizations if they have been used for production of a range of products already; (3) many tools for genome editing are available; and (4) many gene expression tools are available, e.g., plasmids, promoters, and terminators. Each of the above mentioned cell factories have specific advantages. For example, *A. niger* and *B. subtilis* have very efficient protein secretion and are therefore widely used for production of industrial enzymes, while CHO cells are well suited for production of glycosylated proteins to be used as pharmaceuticals. For fuels and chemicals, there is an increasing focus on use of *S. cerevisiae* and *E. coli* as platform cell factories, with *C. glutamicum* as an attractive third choice. To produce a molecule of interest, the biosynthetic pathway for the molecule is reconstructed in the platform cell factory, resulting in establishment of a proof-of-principle strain (Figure 1B). Generally, this strain can be patented and represents a key milestone in the development of a novel bioprocess.

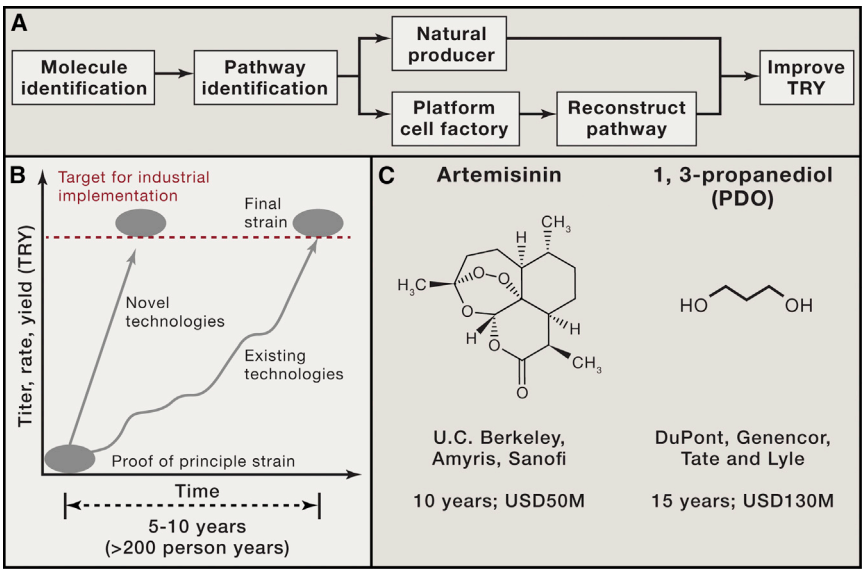

Figure 1. Development of Novel Bioprocesses

(A) The typical workflow for developing a biotech process for production of a new molecule. TRY stands for titer, rate, and yield.

(B) With current technologies, development of a final strain that can be used for industrial production from a proof-of-principle strain takes several years and is costly. There is a need for new technologies that can shorten the development time and reduce the costs.

(C) Example of time and cost for development of bioprocesses for two molecules that have been launched on the market, the anti-malarial drug artemisinin and the chemical building block 1,3 propanediol.

Improving Strain Performance

The road from development of a proof-of-principle strain to having a cell factory that can be used for commercial production is long and arduous. The majority of operational costs come with the fermentation process (Caspeta and Nielsen, 2013), primarily due to relatively high feedstock costs, and being cost-competitive therefore translates to specific demands on titer (final concentration in the fermentation medium), rate (production per unit of time), and yield (units of product synthesized per unit of raw material consumed), often referred to as titer, rate, and yield (TRY) requirements. Moving from a proof-of-principle strain to a production strain that meets industrial TRY requirements is the last but most challenging part of developing a novel bioprocess (Figure 1A), typically involving many years of costly development time (Figure 1B and 1C).

The main reason for the long development time is the need to go through many rounds of strain construction and subsequent phenotypic characterization. Most strains used for industrial production require a large number of genetic modifications, not only in the pathways of interest, but also in other pathways in order to efficiently redirect metabolic flux. For example, in the *E. coli* strain used for production of 1,3-propanediol (used for production of polymers and solvents), the phosphotransferase (PTS) transport system for glucose uptake and phosphorylation was replaced by a heterologous glucose transporter and an additional hexokinase (Nakamura and Whited, 2003). This was done in order to decouple glucose transport from the lower glycolysis, making it possible to convert glucose to 1,3-propanediol with higher yield. In *S. cerevisiae,* improved ethanol and reduced glycerol production could be obtained by engineering the glutamate biosynthetic pathway (Nissen et al., 2000). By replacing the NADPH-dependent glutamate dehydrogenase with a NADH-dependent pathway, ammonia uptake became linked to NADH consumption. With this new NADH "sink," glycerol production was reduced, freeing up more carbon for ethanol production. Traditionally, each round of genetic engineering could only be done in a serial fashion, so it was time consuming to introduce the many genetic modifications required for a final production strain. As we will discuss later, a number of new technologies are likely to change this and reduce the time and cost of strain development.

The Bow-Tie Structure of Metabolism

There is a fundamental biological reason why it is often necessary to make a large number of genetic modifications to alter cell metabolism. Metabolism is one of the conserved features of all living cells and has evolved to be organized into a "bow-tie" structure (Figure 2A). This means that all carbon and energy sources are converted through central carbon metabolism pathways into a set of 12 precursor metabolites (Figure 2A) that are used for biosynthesis of all cellular components and natural products generated by cells (Neidhardt et al., 1990). This results in high flux of carbon through most of the precursor metabolites, each of which is involved in a large number of reactions (Nielsen, 2003). For example, in yeast, acetyl-CoA is involved in 34 compartmentalized metabolic reactions, besides being used for acetylation of macromolecules. To balance the use of these precursor metabolites, cells have evolved several levels of tight regulation, especially to control biosynthesis of amino acids, lipids, nucleotides, and carbohydrates needed for cell growth, homeostasis, and maintenance. It is due to this tight regulation that redirecting the carbon fluxes in central carbon metabolism toward molecules of interest is inherently so difficult.

Regulation of central carbon metabolism has evolved to ensure that production of cellular components is balanced with energy production and consumption. This allows cells to maintain metabolic homeostasis even when exposed to varying environmental and nutritional conditions. The same biological and thermodynamic principles that allow cells to be robust and maintain homeostasis also make metabolic engineering challenging. On the other hand, this robustness can be an advantage. Indeed, many industrial processes take advantage of cells' ability to maintain homeostasis in changing and often harsh industrial conditions, such as stress imposed by high osmolality, varying temperatures, low pH, and high product concentrations that are often toxic. For these reasons, industry often prefers robust

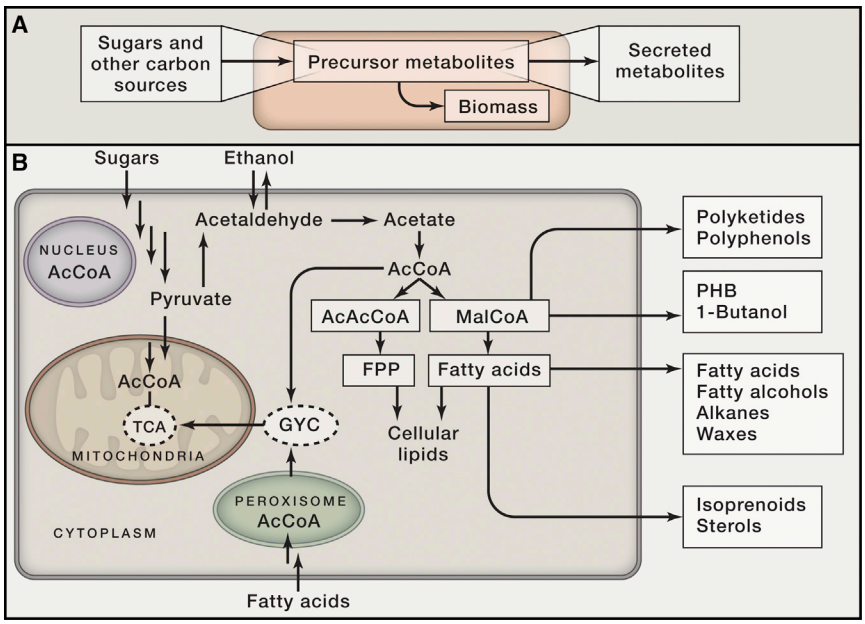

Figure 2. The Bow-Tie Structure of Metabolism and Acetyl-CoA Metabolism in Yeast

(A) According to the bow-tie structure of metabolism, all carbon sources are converted to 12 precursor metabolites that are used for biosynthesis of all secreted metabolites. The precursor metabolites are also used for the biosynthesis of all building blocks that are needed for synthesizing macromolecules making up the biomass of the cell. The 12 precursor metabolites are: glucose-6-phosphate, fructose-6-phosphate, ribose-5-phosphate, erythrose-4-phosphae, glyceraldehyde-3-phosphate, 3-phosphoglycerate, phosphoenol-pyruvate, pyruvate, acetyl-CoA, 2-oxoglutarate, succinyl-CoA, and oxaloacetate.

(B) Illustration of how an acetyl-CoA over-producing strain can be used as a platform strain for production of a range of different molecules. Acetyl-CoA (AcCoA) metabolism in yeast is compartmentalized and there is no direct exchange of this metabolite between the different compartments. AcCoA is formed in the mitochondria from pyruvate and enters the tricarboxylic acid cycle (TCA). AcCoA is also formed in the peroxisome from either fatty acids or acetate and can, via the glyoxylate cycle (GYC), be converted to malate that can be transported to the mitochondria for oxidation. In order to ensure efficient secretion of the product from the cell it is generally preferred to reconstruct the heterologous pathway in the cytosol, and there is therefore a need to ensure efficient provision of cytosolic AcCoA. AcCoA in the cytosol is produced from acetate and is used for production of acetoacetyl-CoA (AcAcCoA), required for the biosynthesis of sterols via farnesyl pyrophosphate (FPP), and for production of malonyl-CoA (MalCoA), required for fatty acid biosynthesis. AcAcCoA, MalCoA, FPP, and fatty acids can all be converted to commercially interesting products.

cell factories that not only survive, but also divide and produce the product of interest even under such adverse conditions.

The yeast *S. cerevisiae* has a proven record of large-scale production of bioethanol and is a favorite organism within industry, but its central carbon metabolism is extensively regulated and has a relatively "flat" structure, with transcriptional regulation alone involving 102 transcription factors (TFs), 78% of which are connected by cross-regulation in a large internal regulatory loop (Österlund et al., 2015). Like most bacteria, *E. coli* has a more hierarchical TF network structure (Yu and Gerstein, 2006), making it easier to redirect carbon fluxes to overproduce a specific molecule (Chen et al., 2013a), with two prominent examples being 1,4-butanediol (Yim et al., 2011) and short alkanes (Choi and Lee, 2013). In addition, several recent studies in *E. coli* have provided detailed new knowledge of metabolic regulation, such as control of iron metabolism through the Fur transcriptional regulatory network (Seo et al., 2014) and mechanisms of oxidative stress metabolism (Seo et al., 2015). Such insights will allow for improved design and faster development of cell factories.

Principles and Tools for Advancing Metabolic Engineering

Platform Strains

Even though the bow-tie structure of metabolism is a challenge for metabolic engineering, it also offers some features that may accelerate strain development in the future. For instance, imagine that for one project, a strain is developed to convert a carbon source (e.g., glucose) into a molecule of interest by efficiently funneling it through an intermediate molecule (e.g., acetyl-CoA) at the center of the bow-tie. With additional smaller modifications, this strain could then become a platform for creating other strains to synthesize products derived from that same intermediate. Since the hardest problem in strain development is often deregulation of central carbon metabolism, such a strain would be of great value, as the development of the new strain from that step onward would proceed relatively fast.

This concept of platform strains (Nielsen, 2015) is by no means new and has been applied successfully before. For example, the Dutch company DSM, the largest producer of β-lactam antibiotics in the world, used one of their high-yielding penicillin-producing strains as a platform strain to engineer the fungus *P. chrysogenum* to efficiently produce 7-ADCA, from which cephalosporins can be derived. They achieved this by extending the penicillin biosynthetic pathway with an expandase, combined with feeding the cells adipic acid (Crawford et al., 1995), thereby leveraging the many years of work that went into developing efficient penicillin-producing strains to generate a new and more valuable product. Similarly, the Danish company Novozymes, the largest enzyme producer in the world, has used strains of the fungus *Aspergillus oryzae* that have been optimized for protein secretion to rapidly develop efficient production processes for new fungal enzymes to be used in detergents, the food industry, and the biofuel industry.

Platform strains were also used early on in the development of *E. coli* strains that efficiently produce aromatics. Bio-based production of aromatics has attracted much interest from the chemical industry, as many molecules of industrial value, such as aspartame and indigo, can be derived from aromatic amino acids or their intermediates. Reconstruction of the *E. coli* pathway for conversion of the amino acid tryptophan into the plant-derived dye indigo represented a key milestone in metabolic engineering (Murdock et al., 1993). Following this, there

were several successful cases of engineering *E. coli* metabolism to overproduce aromatics. In one study, Liao and colleagues increased the supply of phosphoenolpyruvate (PEP), a precursor metabolite for biosynthesis of aromatics, by either expressing a PEP synthase (Patnaik and Liao, 1994) or using a non-PTS sugar transport system (Patnaik et al., 1995).

Recently, *S. cerevisiae* has also been engineered for high-level production of aromatics (Rodriguez et al., 2015), with the objective of producing natural plant products such as stilbenoids and flavonoids. In these cases, one can take advantage of prior knowledge from plant engineering, since it is generally relatively easy to express plant P450 enzymes in *S. cerevisiae*. Indeed, there are numerous examples of reconstructing complex plant pathways using aromatic amino acids as building blocks in yeast. These include production of the antioxidant and potential drug resveratrol, which is found in the skin of grapes (Li et al., 2015), and the antioxidant naringenin, which has anti-inflammatory and immune-stimulating effects (Koopman et al., 2012). Notably, the recent reconstruction of a 23-enzyme pathway to produce opioids in yeast (Galanie et al., 2015) represents an important milestone in the field, as it shows that even very long and complex pathways can be successfully reconstructed. This study illustrates another advantage of using a platform cell factory: having a strain with increased flux toward tyrosine, the precursor for the biosynthetic pathway, made it easier to identify good candidate enzymes for the pathway (Galanie et al., 2015). Despite the success, however, obtaining a proof-of-principle strain producing a low titer of the product is only the first step toward establishing a commercial process, and the TRY of opioid production will need to be significantly improved before microbial production can replace the current process with extraction from plants.

One area that has attracted significant attention recently is the development of yeast platform strains to produce acetyl-CoA, as many chemicals of interest can be derived from this precursor metabolite (Nielsen, 2014; Krivoruchko et al., 2015). Many commodity chemicals and advanced biofuels must be produced in large quantities, and using yeast as a cell factory is therefore favorable, as current bioethanol plants could be retrofitted to produce these more valuable chemicals. However, as illustrated in Figure 2, acetyl-CoA metabolism in yeast is compartmentalized. In the cytosol, acetyl-CoA is used for lipid biosynthesis, either via malonyl-CoA for fatty acids or acetoacetyl-CoA for sterols via the mevalonate pathway, and is derived from acetate by acetyl-CoA synthetases (Acs). Acetate comes from acetaldehyde, an intermediate in the conversion of pyruvate to ethanol, the key fermentative route for yeast. On the other hand, acetyl-CoA in the mitochondria is formed from pyruvate by the pyruvate dehydrogenase (Pdh) complex, and there is no direct exchange of acetyl-CoA between the two compartments, although acetyl-CoA in the cytosol can be transported to the mitochondria via malate or succinate (Chen et al., 2012a). Even though biosynthetic pathways can be reconstructed in the mitochondria (Avalos et al., 2013), it is generally preferable to do so in the cytosol, as this facilitates export of the final product, which in turn facilitates isolation and purification of the desired compound and reduces the production costs dramatically.

The biosynthesis of lipids is highly regulated, particularly at two enzymatic steps, the conversion of acetyl-CoA to malonyl-CoA by acetyl-CoA carboxylase (Acc) and the conversion of hydroxy-3-methyl-glutaryl-CoA (HMG-CoA) into mevalonate by HMG-CoA reductase (Hmgr). Acc is inactivated at the protein level by the protein kinase Snf1 (AMPK in human) (Nielsen, 2009), a global energy regulator (Usaite et al., 2009). Recently, it was shown that a mutant version of Acc that cannot be phosphorylated enables high flux toward malonyl-CoA (Shi et al., 2014). Hmgr is also regulated at the protein level and is bound to the endoplasmic reticulum membrane while facing the cytosol. By sensing ER membrane sterol composition, Hmgr is feedback inhibited by the presence of ergosterol (Nielsen, 2009). Several studies have shown that flux toward mevalonate can be increased significantly through deregulation of Hmgr by deleting its membrane-binding domain (Donald et al., 1997).

The Acs enzyme is also believed to be regulated through phosphorylation and acetylation, but the exact sites have not been identified. A breakthrough in increasing flux toward acetyl-CoA-derived products was the expression of a mutant version of Acs from *Streptococcus enterica* that carries a point mutation preventing inactivation by phosphorylation (Shiba et al., 2007). Expression of this heterologous Acs is often combined with overexpression of *ALD6* (Shiba et al., 2007), which catalyzes the conversion of acetaldehyde to acetate. This strategy was recently combined with blocking of the glyoxylate cycle to prevent transfer of acetyl-CoA from the cytosol to the mitochondria (Chen et al., 2013b). However, the Acs-catalyzed reaction involves conversion of ATP to AMP, so several studies have aimed at creating an energetically more efficient pathway from pyruvate to acetyl-CoA in the cytosol. For example, some groups have heterologously expressed bacterial pyruvate formate lyase, which converts pyruvate to formate and acetyl-CoA (Waks and Silver, 2009; Kozak et al., 2014a; Zhang et al., 2015), where formate can subsequently be oxidized to carbon dioxide, with the generation of NADH, by formate dehydrogenase. Alternatively, a bacterial Pdh localized to the cytosol can directly generate acetyl-CoA from pyruvate (Kozak et al., 2014b), but this is a major undertaking as this enzyme is a multimeric and is larger than bacterial ribosomes.

These studies teach the general lesson that it is often necessary to combine overexpression of specific enzymes with deregulation of the pathway in order to ensure high flux through the pathway of interest. An alternative to de-regulation of individual enzymes is the expression of a complete heterologous pathway, as illustrated by expression of the yeast mevalonate pathway in *E. coli* (Martin et al., 2003). *E. coli* uses a non-mevalonate pathway for the biosynthesis of farnesyl pyrophosphate, an intermediate of the sterol biosynthetic pathway and a precursor for biosynthesis of sesquiterpenes, a broad class of chemicals that can be used as perfumes, pharmaceuticals, and biofuels. This approach circumvents the problem of the endogenous pathway being regulated, resulting in a significant increase in flux toward farnesyl pyrophosphate, an intermediate for the anti-malarial drug artemisinic acid (Martin et al., 2003).

Genetic Tools

One of the key requirements for metabolic engineering is the availability of good genetic tools for genetic engineering of the

host cell (Redden et al., 2014; Jensen and Keasling, 2014; David and Siewers, 2014). As mentioned above, manipulation of metabolism generally involves the knockout, introduction, and overexpression or mutation of more than one gene. Although using autonomously replicating vectors, such as plasmids, to introduce genes is useful for constructing proof-of-principle strains, plasmids tend to be unstable when used in large-scale industrial cultivation that involves massive cell expansion. In the past, introducing genes into chromosomes was accomplished primarily using phage integration sites in bacteria and homologous recombination in yeast. However, "clustered regularly interspaced short palindromic repeats" (CRISPR)/CRISPR-associated protein Cas9-based systems now allow introduction of genes into nearly any location in the chromosome (Jinek et al., 2012; Jakočiūnas et al., 2015). With the ability to vary promoter (Jensen and Hammer, 1998; Redden and Alper, 2015) and ribosome binding strength (Salis et al., 2009), as well as the stability of the mRNA (Smolke et al., 2000; Pfleger et al., 2006) and the resulting protein, there are many levers other than copy number that can be used to alter enzyme production. Morever, in cases where copy number limits protein production, one can amplify genes on the chromosome to increase copy number (Tyo et al., 2009).

Promoters play an essential role in controlling biosynthetic pathways. Inducible promoters are often essential for pathways that produce toxic products, and several inducible expression systems are now available for use in bacteria, yeasts, and other organisms (Wang et al., 2012). Ensuring that these promoters have consistent, tunable control in all cells in a culture is essential for consistent production of the desired molecule and for preventing non-producer cells from taking over the population (Khlebnikov et al., 2001; Lee and Keasling, 2005). Promoters that are constitutive, induced by starvation or upon entry into stationary phase, or quorum-sensing allow for inexpensive, inducer-free gene expression, which is particularly important in large-scale production of chemicals and fuels, where the cost of the inducer is an issue (Tsao et al., 2010). However, a trade-off with using constitutive expression of pathway enzymes is that these often may account for a major fraction of the cellular proteome. Although small non-coding RNAs can be used to control protein expression (Na et al., 2013), so far there have been relatively few implementations of this approach.

Production of most molecules of interest often requires several enzymes, and the expression of the genes encoding these enzymes must be coordinated. There are many ways to coordinate expression of multiple genes: (1) use different inducible promoters for each gene; (2) use the same inducible promoter for each gene but vary the promoter strength (Bakke et al., 2009); (3) use a non-native RNA polymerase or transcription factor to control the expression of more than one gene (Alper and Stephanopoulos, 2007); (4) group multiple, related genes into operons (and use internal ribosomal entry sequences in eukaryotes (Komar and Hatzoglou, 2005); (5) vary the ribosome binding strength for the enzymes encoded in the operon (Salis et al., 2009); (6) control segmental mRNA stability of each coding region (Smolke et al., 2000; Pfleger et al., 2006); (7) control the stability of each enzyme, and (8) spatial control through attachment to a protein scaffold (Dueber et al., 2009) or targeting to special organelles (Farhi et al., 2011; Avalos et al., 2013).

In all of these cases, it is desirable for the metabolic engineer to know the specific activity of each enzyme in the pathway in order to design promoter or ribosome binding site strength or the stability of mRNA or protein in order to "dial in" the correct amount of enzyme in the pathway. However, as knowledge about the activity of each enzyme in vivo is often absent, the levels of each metabolite and enzyme in the pathway must be measured to determine if there are any pathway bottlenecks and then the level of expression (or mRNA or protein stability) of the limiting enzyme must be adjusted. This can be a laborious process. The development of dynamic regulators using transcription factors that can sense intermediates in the biosynthetic pathway (Farmer and Liao, 2000; Zhang et al., 2012) or promoters that respond to stress (Dahl et al., 2013) eliminates the need to regulate every step of the pathway and puts the control in the hands of the cell. Similarly, gene expression can be controlled in response to medium components, as illustrated by promoters for hexose transporters in yeast allowing dynamic regulation of gene expression in response to the extracellular glucose concentration, which can be used to downregulate a pathway competing for the precursor needed for the desired product (Scalcinati et al., 2012).

Regardless of how sophisticated the design tools and how good the blueprint, there will always be "bugs" in the engineered system, as we do not know everything about how metabolism is regulated. For the development of microbial cell factories, systems biology can provide debugging routines (Park et al., 2007; Park et al., 2014; Caspeta et al., 2014; Kizer et al., 2008). Through transcriptomic, proteomic, and metabolomic measurements combined with integrative analysis, it is possible to get insight into how the introduction of a metabolic pathway impacts overall cellular physiology. Often, expression of a heterologous metabolic pathway elicits a stress response in the host, due to protein overproduction or accumulation of toxic intermediates or end products (Gill et al., 2000; Martin et al., 2003). These stresses are reflected in mRNA and protein expression and can therefore be identified using analysis of the transcriptome, proteome, metabolome, or fluxome. Information from one or more of these techniques can then be used to modify expression of genes in the metabolic pathway or in the host to improve titers and/or productivity of the final product.

Adaptive Laboratory Evolution and High-Throughput Screening

Once an organism is constructed with a desired metabolic pathway, it is necessary to further optimize the metabolic pathway to increase TRY. Besides directed modification of gene expression, as described above, TRY can be improved using adaptive laboratory evolution (ALE) (Dragosits and Mattanovich, 2013). If production of the desired chemical is coupled to growth (that is, when the cells grow they must produce the chemical), then one can use improvements in the growth of the organism to improve the production of the desired molecule. ALE is one way to select for faster growing organisms, thereby selecting for higher production of the desired molecule, as illustrated by succinic acid production by yeast (Otero et al., 2013). In this study, the normal route for biosynthesis of glycine was

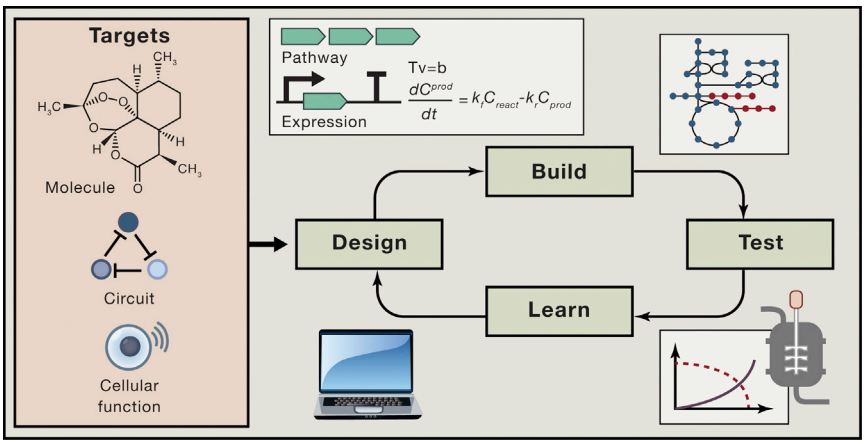

Figure 3. The Design-Build-Test-Learn Cycle of Metabolic Engineering

Following identification of a target molecule, a regulatory circuit to be used for expression, and a suitable host, the biological systems are designed. This may involve the use of mathematical models of metabolism and BioCAD software designing optimal constructs. Thereafter, the pathway is reconstructed in the build phase and the central carbon metabolism is engineered to ensure efficient provision of the precursor metabolite. The constructed strain is tested in bioreactors that simulate industrial-like conditions, and following analysis of the data, new knowledge is generated. This is stored in the learn phase of the cycle and can hereby be used for improved design in the next round.

deleted, and an alternative route was introduced that resulted in production of succinic acid as a by-product, so succinic acid became a growth-coupled metabolite. ALE has also been shown to be very efficient for improving growth on non-preferred carbon sources, such as glycerol for *E. coli* (Ibarra et al., 2002), galactose for yeast (Hong et al., 2011), and xylose for yeast (Kuyper et al., 2005), as well as for improving the tolerance to harsh conditions or to the product of interest, as reviewed recently (Dragosits and Mattanovich, 2013).

Through the use of next-generation sequencing and systems biology, it is possible to identify mutations responsible for the desirable phenoptypes. For example, a single mutation in the *ERG3* gene conferred upon yeast the ability to grow at elevated temperatures (Caspeta et al., 2014). In this study, deep sequencing of the genome gave clear hints on causal mutations, but transcriptome and/or metabolome analysis assisted in mapping molecular mechanisms underlying the acquired phenotypes. Thus, the mutation was found to result in altered sterol composition (ergosterol in the yeast membrane was replaced by fecosterol), and this was associated with an upregulation of sterol metabolism. This showed that altered membrane properties due to changes in sterol composition allowed for improved growth at elevated temperatures.

Although it is trivial to tie substrate consumption or stress tolerance to growth, coupling production of the majority of small molecules of commercial interest—such as fatty acids, diols, diamines, and short-chain alcohols among others—to growth is difficult. It is therefore necessary to use other screening or selection methods to identify improved strains. The combination of microtiter plates for growth of strain libraries with gas and liquid chromatography techniques is an option, but the throughput (10^2–10^3 variants per machine per day) falls far short of levels necessary for effective interrogation of large genetic libraries. Microfluidic cell sorting offers interesting opportunities for screening of cell libraries, as demonstrated recently for identification of yeast strains with improved xylose uptake (Wang et al., 2014), *E. coli* strains with improved lactic acid production (Wang et al., 2014), and yeast strains with improved protein secretion capacity (Huang et al., 2015).

In nature, the need for sensitive and specific small-molecule detection and response has been addressed in part through evolution and selection for ligand-responsive transcription factors and their cognate promoters. Transcription-factor-promoter pairs are archetypal genetic devices within the synthetic biology paradigm. Abundant in nature, highly modular, and capable of being evolved or re-engineered, transcription-factor-based devices are well suited for a broad range of applications. While engineered transcription-factor-based biosensors have been employed for detection of exogenous environmental pollutants (Simpson et al., 1998), this work has only recently been explored in the context of metabolic engineering (Chou and Keasling, 2013). Through coupling increased production of an intracellular metabolite with expression of fluorescent proteins, fluorescent-activated cell sorting (FACS) has been used for screening of strains with improved phenotype. Recently, transcription-factor-based detection of small molecules has been used to increase production of adipate, succinate, and 1-butanol (Dietrich et al., 2013). We anticipate a time when biosensors can be readily made for any desired product, allowing use of high-throughput screening using FACS or microfluidics, and hereby, significantly reduce the time required for improving the TRY.

Design-Build-Test-Learn Cycle

As described above, the typical process for engineering metabolism, as any other system, involves four highly interdependent modules (Figure 3): design (D) of a biological system, in this case metabolic pathways in a microorganism, to produce a desired molecule and coding of these pathways into DNA parts and assembly instructions; build (B) the biological system from DNA parts and production-relevant microbial chassis, using inputs from D and tools developed through synthetic biology; test (T) to determine if and how the engineered biological system from B carried out the desired function, using cell physiology and omics (possibility to integrate via systems biology tools); and learn (L) to glean information from the performance of engineered biosystems to inform decision making in D, B, and T.

Although these steps are now carried out in the research laboratory and a single turn of the DBTL cycle can take months of work (Qin et al., 2015), we envision a time when metabolic engineering will more closely resemble electronics engineering, with turn-around times on the order of days to a couple of weeks. Computer-aided design software for biology will allow the

metabolic engineer to design a metabolic pathway inside an organism of interest, send that design to a biological foundry that would construct the pathway in the organism of interest (Chen et al., 2012b), and within a reasonable time frame, send that engineered organism back to the engineer for scale-up and production. In order for the foundry to be able to reliably construct a functional metabolic pathway inside the target organism, the foundry will need all of the tools to build the pathway (e.g., robotic liquid handling or microfluidics for DNA construction [Shih et al., 2015], genetic control systems to control the genes of the new metabolic pathway [Lee et al., 2011; Lee et al., 2015], tools to knock out competing pathways inside the host organism, etc.), methods and equipment for growing and assaying for the final product, and above all, machine-learning software to gather the successes and failures of the design, build, and test processes and attempt to learn from those to make the design software more capable during the next round. Although it may be some time before metabolic engineering has the rapid turnarounds of electronics engineering, new technologies as discussed below will clearly lead to a significant reduction in the turnaround time in the DBTL cycle.

Design

Current pathway design is often treated as a one-off process, relying heavily on domain expertise with no standardization. The pathway designer generally determines what organism he/she will use for the production process based on the starting materials available (e.g., sucrose from cane, glucose from starch, mixed sugars from cellulosic biomass), the toxicity of the desired product to an organism, and the processing conditions necessary to produce and purify the desired product (e.g., high temperature, low pH, etc). Based on the choice of organism, the metabolic engineer is provided with an available set of intracellular metabolites from which to produce the desired end product. To get from the available starting metabolites inside the cell to the desired product, the metabolic engineer searches for enzymes that could be used in a heterologous metabolic pathway; these enzymes can be found in online databases of pathways, the literature where metabolic pathways of various organisms are described, and genome sequence databases where annotations might indicate reactions that have little to no documentation in a particular host. In cases where no specific enzyme can be identified, one may evolve an enzyme to carry out the desired reaction (Renata et al., 2015) or construct an enzyme de novo (Siegel et al., 2010), which is quite difficult.

The approach described above is difficult to scale, and is often inefficient because there is no ability to reuse parts or data from related designs. For data capture and exchange, there are local successes in the broader community, such as the systems biology community, where standards have been developed for omics data capture, but there is little formalism around genotype specification, strain construction specification, and particularly formal representation of observations about data. Small-scale labs will frequently capture these data on paper or perhaps a spreadsheet in no particular format, making it extremely problematic to apply these results to an open-source production framework.

Here, a BioCAD software providing information about the starting materials available and the desired product would be extremely useful and this kind of software would identify a range of suitable organisms based on substrates available and process conditions necessary to produce and purify the desired product (e.g., low pH, high temperature, etc). After the user selects the organism, the BioCAD software would then identify all possible pathways between available intermediates in the cell and the final product, e.g., using BNICE.ch (Hadadi and Hatzimanikatis, 2015). Furthermore, using detailed metabolic models, BioCAD would be able to enumerate different metabolic engineering targets that would improve the yield in the conversion of the substrate to the product. Here, genome-scale metabolic models (GEMs) have shown to be particularly useful (O'Brien et al., 2015; Lee and Kim, 2015), and GEMs have been developed for most industrially relevant microorganisms (Kim et al., 2012; Garcia-Albornoz and Nielsen, 2013). Recently, these models have been expanded to include many other key cellular processes, such as transcription and translation (O'Brien and Palsson, 2015), allowing for improved simulation capabilities of these models.

A strength of these models is that they are "open-ended," meaning that new information can be added to the models when it is acquired. This was illustrated in a recent study on oxidative stress in *E. coli*, where several key pathways were identified to be missing in the GEM and improved performance of the model when added (Brynildsen et al., 2013). These models do, however, have limitations as they only provide stoichiometric constraints. Thus, efforts have been made toward integrating kinetic information into GEMs, and in this case, the BioCAD software could also have the V_{max} and K_m values for all of the necessary enzymes—as well as dependencies of the enzymes for cofactors, pH, temperature, etc.—so that promoters, mRNA stabilities, and enzyme stabilities could be programmed to deliver the most appropriate enzyme for each step in the correct amount to achieve the desired reaction. Once all design alternatives are evaluated, the best choice would be sent for construction.

Build

The build phase is the construction or retrofitting of the metabolic pathway in the desired host, as well as the deregulating of the central carbon metabolism such that a higher flux can be directed toward the product of interest. Pathway reconstruction includes synthesis of large DNA constructs containing the genes encoding the enzymes of the metabolic pathways and the associated genetic control systems to regulate enzyme production. Build also includes knocking out genes or pathways that might compete or otherwise interfere with the functioning of the heterologous metabolic pathway.

Large DNA construction is one area that has greatly expanded over the past several years (Kosuri and Church, 2014). It is now possible to purchase long DNA that will encode an entire enzyme or a series of enzymes to constitute an entire metabolic pathway. This has greatly reduced the time and effort needed to build metabolic pathways, allowing the metabolic engineer to focus more on developing the host.

In theory, any build team would have at their disposal a variety of host organisms that have different characteristics: different optimal growth temperatures, pH optima, abilities to tolerate various chemicals, abilities to consume different carbon sources, etc. Ideally, these hosts would all be transformable and

have well-characterized genetic systems that would allow for control of transcript and protein abundance and timing of pathway activity during the various phases of growth. In reality, this is rarely the case: only a few hosts are known well enough to allow rapid and easy construction of metabolic pathways, and even in the case of well-known hosts, the genetic tools are rarely characterized to the extent that the desired level of the metabolic pathway can be programmed accurately. For instance, expression from a specific promoter may be context dependent and, therefore, vary depending on what other genetic modifications are introduced into the host cell.

The recent development of CRISPR/Cas9 systems has allowed engineering of nearly any host that is transformable (Jinek et al., 2012). Modifications of these systems allow insertion of many genes into many target sites (Jakočiūnas et al., 2015), knock out or downregulation of competing pathways (Gilbert et al., 2013), and upregulation of beneficial pathways. These methods will likely continue to be used and will become a standard tool in the genetic-engineering toolbox. Furthermore, even though well-characterized promoters, ribosome binding sites, mRNA stability elements, and the like are limited, the development of computer algorithms to calculate native promoter and ribosome binding site strength and then to design new ones will greatly facilitate construction of metabolic pathways that perform as desired (Salis, 2011).

Test

The test phase includes anything that determines the efficacy of the design and build, including but not limited to (1) verification of Build success (i.e., construction of metabolic pathway, knockout of specific genes, integration of genes, etc.), (2) growth and physiological characterization of the engineered cells, and (3) measurement of the transcripts, proteins, and/or final products of the engineered pathway, often at genome scale. It is advantageous to use high-throughput methods, e.g., transcriptomics, proteomics, and metabolomics, as these allow for global analysis of cellular metabolism. It is difficult to analyze multiple data types, but GEMs provide a good scaffold for analysis (Patil and Nielsen, 2005; Usaite et al., 2009). High-throughput analysis allows for measurements of specific pathway protein production (Redding-Johanson et al., 2011), specific metabolite presence or perturbation, or specific gene expression (Regenberg et al., 2006; Boer et al., 2010). However, the technologies were developed for low-throughput research and biomarker identification for small numbers of proteins or metabolites and, when adapted for metabolic engineering applications, they are slow and only allow analysis of a smaller subset of strains. As a result, they cannot be used for routine analysis in the test phase, as this would be too costly and time-consuming. The absence of a comprehensive dataset for each constructed strain severely limits improvement in the success rate of the DBTL cycle, so improved technologies for formalizing data capture, data analysis, and data interpretation are needed.

Learn

Learning is possibly the most weakly supported step in current metabolic engineering practice, yet perhaps the most important to increasing the rate of success. It is typically nonsystematic and lacks statistical rigor, relying on ad hoc observations, literature data, and intuition gathered by individual researchers responsible for the next round of pathway design. Failed experiments are often discarded or inaccessible to data mining and seen as uninformative, with only rare success selectively archived. Nonetheless, it is clear that experienced laboratories can more consistently produce target molecules of interest, suggesting an opportunity to formalize the learning process.

One area where the DBTL cycle may particularly contribute to gaining new biological insight is on how metabolism is regulated. We have extensive information about regulation of metabolism, but this is generally based on studies of one or a few regulatory components. A few systems biological studies have enabled global mapping of key regulatory components, e.g., Snf1 in yeast (Usaite et al., 2009), but it is still a challenge to integrate this information into concrete design strategies. However, by integrating engineering design with available information about regulation, possibly combined with targeted new experiments to identify novel regulatory structures, it may be possible to significantly advance our understanding of how metabolism is regulated at the global level. This may allow for new strategies, as targeting regulation can in some cases be better than simply overexpressing specific pathway enzymes. For example, yeast galactose uptake rate was not improved by overexpression of individual enzymes or combinations of enzymes in the Leloir pathway (de Jongh et al., 2008), whereas a 40% increase could be obtained by engineering the GAL-regulon (Ostergaard et al., 2000)

Recently, statistical techniques such as principal component analysis (PCA) have been used to analyze data from engineered organisms to inform the next round of design (Alonso-Gutierrez et al., 2015). In the past, data from proteomics analysis were too complicated to allow one to deduce trends that could be used to understand system limitations and to reengineer the system, but techniques like PCA allow the analysis of small datasets to reveal patterns or trends that can be used to guide re-design of a biological system. As more data of any one type and more diverse data are collected, it will be necessary to use more sophisticated data-analysis tools, such as machine-learning algorithms (Tarca et al., 2007). Machine-learning techniques are being used in a diverse set of applications, but to date, it has been used relatively little for metabolic engineering purposes. It may, however, offer the possibility of deducing patterns and trends that will aid in redesign of biological systems.

At this time, many biological engineering exercises still do not collect the vast amounts of data that are collected in other settings, such as over the internet. With improvements in the types and speed with which we can collect data on engineered systems, we will soon be awash in data and will need computational methods to make sense of it all. This will allow us to identify bottlenecks in biosynthetic pathways, diagnose exactly why the bottlenecks exist, and reengineer systems to produce higher TRYs of the desired product in less time with less human intervention.

Perspectives

The development of cell factories, which can be used for cost-efficient production of fuels, chemicals, foods, feeds, and pharmaceuticals, requires multiple rounds of the DBTL cycle, often

because we are missing knowledge of how metabolism is regulated. This takes time and is costly. The main reason for this is the extensive robustness of cell metabolism, which is due to redundancy, regulation, and tight interaction of metabolism and all other cellular processes. Metabolism has evolved to support cell growth and maintenance, and when we seek to engineer metabolism to redirect metabolic fluxes toward a specific metabolite, the regulation within the cell will strive to keep homeostasis and, therefore, counteract our engineering efforts. However, by formalizing the learn part of the DBTL cycle, it will be possible to capture knowledge generated in different metabolic engineering efforts and, thereby, accelerate the process. This will require establishment of BioCAD software that can integrate knowledge and be used as an interactive tool for improved design by the metabolic engineer. We envisage that BioCAD can also hold information about detailed metabolic models for platform cell factories and information about promoters, terminators, integration sites, vectors, and more, so that the complete design process can be automated. BioCAD could also be used to integrate so-called ''big data'', e.g., where multi-omics data from many different strains are collected and analyzed in an integrative fashion. Together with information about transcription-factor networks and protein-protein interaction networks, this could be used to gain much new insight into regulation of the applied cell factory. This will allow the metabolic engineer to rapidly test different designs and score these against each other and, thereby, facilitate the design phase. With the advancement in DNA synthesis and robotics for cloning and phenotypic characterization, the build and test processes may also, to a large extent, be automated, and the development of novel cell factories will develop similarly as seen in other manufacturing processes.

Even though we do have extensive knowledge about yeast and *E. coli* that can be integrated into a future BioCAD, a major hindrance for advancement of the field is our lack of fundamental knowledge. We mentioned several of these earlier, but it will also be necessary to expand our current list of platform cell factories in order to expand the possibilities for biochemical transformations. Not all pathways express well in yeast and *E. coli,* and it may also be necessary to have cell factory platforms that can operate at extreme temperatures, extreme pH values (high and low), and extreme salt concentrations. The development of a solid knowledge base for such new platform cell factories will obviously be time-consuming, but using the scaffold for knowledge integration established through BioCAD, it will be possible to advance rapidly.

ACKNOWLEDGMENTS

We acknowledge funding from the Novo Nordisk Foundation, the Knut and Alice Wallenberg Foundation, Vetenskapsrådet, FORMAS, the Joint BioEnergy Institute (http://www.jbei.org/), which is supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the US Department of Energy, and by the Synthetic Biology Engineering Research Center (SynBERC) through National Science Foundation grant NSF EEC 0540879. We thank Yun Chen, Victor Chubukov, Nathan Hillson, Mingtao Huang, Dina Petranovic, and Yongjin Zhou for constructive comments.

REFERENCES

Ajikumar, P.K., Xiao, W.-H., Tyo, K.E., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T.H., Pfeifer, B., and Stephanopoulos, G. (2010). Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. Science *330*, 70–74.

Alonso-Gutierrez, J., Kim, E.-M., Batth, T.S., Cho, N., Hu, Q., Chan, L.J.G., Petzold, C.J., Hillson, N.J., Adams, P.D., Keasling, J.D., et al. (2015). Principal component analysis of proteomics (PCAP) as a tool to direct metabolic engineering. Metab. Eng. *28*, 123–133.

Alper, H., and Stephanopoulos, G. (2007). Global transcription machinery engineering: a new approach for improving cellular phenotype. Metab. Eng. *9*, 258–267.

Avalos, J.L., Fink, G.R., and Stephanopoulos, G. (2013). Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. Nat. Biotechnol. *31*, 335–341.

Bailey, J.E. (1991). Toward a science of metabolic engineering. Science *252*, 1668–1675.

Bakke, I., Berg, L., Aune, T.E., Brautaset, T., Sletta, H., Tøndervik, A., and Valla, S. (2009). Random mutagenesis of the PM promoter as a powerful strategy for improvement of recombinant-gene expression. Appl. Environ. Microbiol. *75*, 2002–2011.

Boer, V.M., Crutchfield, C.A., Bradley, P.H., Botstein, D., and Rabinowitz, J.D. (2010). Growth-limiting intracellular metabolites in yeast growing under diverse nutrient limitations. Mol. Biol. Cell *21*, 198–211.

Brynildsen, M.P., Winkler, J.A., Spina, C.S., MacDonald, I.C., and Collins, J.J. (2013). Potentiating antibacterial activity by predictably enhancing endogenous microbial ROS production. Nat. Biotechnol. *31*, 160–165.

Caspeta, L., and Nielsen, J. (2013). Economic and environmental impacts of microbial biodiesel. Nat. Biotechnol. *31*, 789–793.

Caspeta, L., Chen, Y., Ghiaci, P., Feizi, A., Buskov, S., Hallström, B.M., Petranovic, D., and Nielsen, J. (2014). Biofuels. Altered sterol composition renders yeast thermotolerant. Science *346*, 75–78.

Chen, Y., Siewers, V., and Nielsen, J. (2012a). Profiling of cytosolic and peroxisomal acetyl-CoA metabolism in *Saccharomyces cerevisiae*. PLoS ONE *7*, e42475.

Chen, J., Densmore, D., Ham, T.S., Keasling, J.D., and Hillson, N.J. (2012b). DeviceEditor visual biological CAD canvas. J. Biol. Eng. *6*, 1.

Chen, X., Zhou, L., Tian, K., Kumar, A., Singh, S., Prior, B.A., and Wang, Z. (2013a). Metabolic engineering of *Escherichia coli*: a sustainable industrial platform for bio-based chemical production. Biotechnol. Adv. *31*, 1200–1223.

Chen, Y., Daviet, L., Schalk, M., Siewers, V., and Nielsen, J. (2013b). Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. Metab. Eng. *15*, 48–54.

Choi, Y.J., and Lee, S.Y. (2013). Microbial production of short-chain alkanes. Nature *502*, 571–574.

Chou, H.H., and Keasling, J.D. (2013). Programming adaptive control to evolve increased metabolite production. Nat. Commun. *4*, 2595.

Crawford, L., Stepan, A.M., McAda, P.C., Rambosek, J.A., Conder, M.J., Vinci, V.A., and Reeves, C.D. (1995). Production of cephalosporin intermediates by feeding adipic acid to recombinant *Penicillium chrysogenum* strains expressing ring expansion activity. Biotechnology (N. Y.) *13*, 58–62.

Dahl, R.H., Zhang, F., Alonso-Gutierrez, J., Baidoo, E., Batth, T.S., Redding-Johanson, A.M., Petzold, C.J., Mukhopadhyay, A., Lee, T.S., Adams, P.D., and Keasling, J.D. (2013). Engineering dynamic pathway regulation using stress-response promoters. Nat. Biotechnol. *31*, 1039–1046.

David, F., and Siewers, V. (2014). Advances in yeast genome engineering. FEMS Yeast Res. *15*, 1–14.

de Jongh, W.A., Bro, C., Ostergaard, S., Regenberg, B., Olsson, L., and Nielsen, J. (2008). The roles of galactitol, galactose-1-phosphate, and phosphoglucomutase in galactose-induced toxicity in Saccharomyces cerevisiae. Biotechnol. Bioeng. *101*, 317–326.

Dietrich, J.A., Shis, D.L., Alikhani, A., and Keasling, J.D. (2013). Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis. ACS Synth. Biol. *2*, 47–58.

Donald, K.A., Hampton, R.Y., and Fritz, I.B. (1997). Effects of overproduction of the catalytic domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase on squalene synthesis in Saccharomyces cerevisiae. Appl. Environ. Microbiol. *63*, 3341–3344.

Dragosits, M., and Mattanovich, D. (2013). Adaptive laboratory evolution – principles and applications for biotechnology. Microb. Cell Fact. *12*, 64.

Dueber, J.E., Wu, G.C., Malmirchegini, G.R., Moon, T.S., Petzold, C.J., Ullal, A.V., Prather, K.L., and Keasling, J.D. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. Nat. Biotechnol. *27*, 753–759.

Farhi, M., Marhevka, E., Masci, T., Marcos, E., Eyal, Y., Ovadis, M., Abeliovich, H., and Vainstein, A. (2011). Harnessing yeast subcellular compartments for the production of plant terpenoids. Metab. Eng. *13*, 474–481.

Farmer, W.R., and Liao, J.C. (2000). Improving lycopene production in *Escherichia coli* by engineering metabolic control. Nat. Biotechnol. *18*, 533–537.

Galanie, S., Thodey, K., Trenchard, I.J., Filsinger Interrante, M., and Smolke, C.D. (2015). Complete biosynthesis of opioids in yeast. Science *349*, 1095–1100.

Garcia-Albornoz, M.A., and Nielsen, J. (2013). Application of genome-scale metabolic models in metabolic engineering. J. Ind. Microbiol. Biotechnol. *9*, 203–214.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442–451.

Gill, R.T., Valdes, J.J., and Bentley, W.E. (2000). A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in *Escherichia coli*. Metab. Eng. *2*, 178–189.

Hadadi, N., and Hatzimanikatis, V. (2015). Design of computational retrobiosynthesis tools for the design of de novo synthetic pathways. Curr. Opin. Chem. Biol. *28*, 99–104.

Hong, K.-K., Vongsangnak, W., Vemuri, G.N., and Nielsen, J. (2011). Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. Proc. Natl. Acad. Sci. USA *108*, 12179–12184.

Huang, M., Bai, Y., Sjostrom, S.L., Hallström, B.M., Liu, Z., Petranovic, D., Uhlén, M., Joensson, H.N., Andersson-Svahn, H., and Nielsen, J. (2015). Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. Proc. Natl. Acad. Sci. USA *112*, E4689–E4696.

Ibarra, R.U., Edwards, J.S., and Palsson, B.O. (2002). Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. Nature *420*, 186–189.

Jakočiūnas, T., Bonde, I., Herrgård, M., Harrison, S.J., Kristensen, M., Pedersen, L.E., Jensen, M.K., and Keasling, J.D. (2015). Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. Metab. Eng. *28*, 213–222.

Jensen, P.R., and Hammer, K. (1998). The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. Appl. Environ. Microbiol. *64*, 82–87.

Jensen, M.K., and Keasling, J.D. (2014). Recent applications of synthetic biology tools for yeast metabolic engineering. FEMS Yeast Res. *15*, 1–10.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Keasling, J.D. (2010). Manufacturing molecules through metabolic engineering. Science *330*, 1355–1358.

Khlebnikov, A., Datsenko, K.A., Skaug, T., Wanner, B.L., and Keasling, J.D. (2001). Homogeneous expression of the P(BAD) promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity AraE transporter. Microbiology *147*, 3241–3247.

Kim, T.Y., Sohn, S.B., Kim, Y.B., Kim, W.J., and Lee, S.Y. (2012). Recent advances in reconstruction and applications of genome-scale metabolic models. Curr. Opin. Biotechnol. *23*, 617–623.

Kizer, L., Pitera, D.J., Pfleger, B.F., and Keasling, J.D. (2008). Application of functional genomics to pathway optimization for increased isoprenoid production. Appl. Environ. Microbiol. *74*, 3229–3241.

Komar, A.A., and Hatzoglou, M. (2005). Internal ribosome entry sites in cellular mRNAs: mystery of their existence. J. Biol. Chem. *280*, 23425–23428.

Koopman, F., Beekwilder, J., Crimi, B., van Houwelingen, A., Hall, R.D., Bosch, D., van Maris, A.J.A., Pronk, J.T., and Daran, J.-M. (2012). De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. Microb. Cell Fact. *11*, 155.

Kosuri, S., and Church, G.M. (2014). Large-scale de novo DNA synthesis: technologies and applications. Nat. Methods *11*, 499–507.

Kozak, B.U., van Rossum, H.M., Benjamin, K.R., Wu, L., Daran, J.M., Pronk, J.T., and van Maris, A.J. (2014a). Replacement of the *Saccharomyces cerevisiae* acetyl-CoA synthetases by alternative pathways for cytosolic acetyl-CoA synthesis. Metab. Eng. *21*, 46–59.

Kozak, B.U., van Rossum, H.M., Luttik, M.A.H., Akeroyd, M., Benjamin, K.R., Wu, L., de Vries, S., Daran, J.-M., Pronk, J.T., and van Maris, A.J.A. (2014b). Engineering acetyl coenzyme A supply: functional expression of a bacterial pyruvate dehydrogenase complex in the cytosol of *Saccharomyces cerevisiae*. MBio *5*, e01696–e14.

Krivoruchko, A., Zhang, Y., Siewers, V., Chen, Y., and Nielsen, J. (2015). Microbial acetyl-CoA metabolism and metabolic engineering. Metab. Eng. *28*, 28–42.

Kuyper, M., Toirkens, M.J., Diderich, J.A., Winkler, A.A., van Dijken, J.P., and Pronk, J.T. (2005). Evolutionary engineering of mixed-sugar utilization by a xylose-fermenting *Saccharomyces cerevisiae* strain. FEMS Yeast Res. *5*, 925–934.

Langer, E.S. (2012). Biomanufacturing outsourcing outlook. Biopharm International *25*. http://www.biopharminternational.com/biomanufacturing-outsourcing-outlook.

Lee, S.K., and Keasling, J.D. (2005). A propionate-inducible expression system for enteric bacteria. Appl. Environ. Microbiol. *71*, 6856–6862.

Lee, S.Y., and Kim, H.U. (2015). Systems strategies for developing industrial microbial strains. Nat. Biotechnol. *33*, 1061–1072.

Lee, T.S., Krupa, R.A., Zhang, F., Hajimorad, M., Holtz, W.J., Prasad, N., Lee, S.K., and Keasling, J.D. (2011). BglBrick vectors and datasheets: A synthetic biology platform for gene expression. J. Biol. Eng. *5*, 12.

Lee, M.E., DeLoache, W.C., Cervantes, B., and Dueber, J.E. (2015). A highly characterized yeast toolkit for modular, multipart assembly. ACS Synth. Biol. *4*, 975–986.

Li, M., Kildegaard, K.R., Chen, Y., Rodriguez, A., Borodina, I., and Nielsen, J. (2015). *De novo* production of resveratrol from glucose or ethanol by engineered *Saccharomyces cerevisiae*. Metab. Eng. *32*, 1–11.

Martin, V.J.J., Pitera, D.J., Withers, S.T., Newman, J.D., and Keasling, J.D. (2003). Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. Nat. Biotechnol. *21*, 796–802.

Murdock, D., Ensley, B.D., Serdar, C., and Thalen, M. (1993). Construction of metabolic operons catalyzing the de novo biosynthesis of indigo in *Escherichia coli*. Biotechnology (N. Y.) *11*, 381–386.

Na, D., Yoo, S.M., Chung, H., Park, H., Park, J.H., and Lee, S.Y. (2013). Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. Nat. Biotechnol. *31*, 170–174.

Nakamura, C.E., and Whited, G.M. (2003). Metabolic engineering for the microbial production of 1,3-propanediol. Curr. Opin. Biotechnol. *14*, 454–459.

Neidhardt, F.C., Ingraham, J.L., and Schaechter, M. (1990). Physiology of the Bacterial Cell (Sunderland, MA: Sinauer Associates).

Nielsen, J. (2001). Metabolic engineering. Appl. Microbiol. Biotechnol. *55*, 263–283.

Nielsen, J. (2003). It is all about metabolic fluxes. J. Bacteriol. *185*, 7031–7035.

Nielsen, J. (2009). Systems biology of lipid metabolism: from yeast to human. FEBS Lett. *583*, 3905–3913.

Nielsen, J. (2014). Synthetic biology for engineering acetyl coenzyme A metabolism in yeast. MBio *5*, e02153.

Nielsen, J. (2015). BIOENGINEERING. Yeast cell factories on the horizon. Science *349*, 1050–1051.

Nissen, T.L., Kielland-Brandt, M.C., Nielsen, J., and Villadsen, J. (2000). Optimization of ethanol production in Saccharomyces cerevisiae by metabolic engineering of the ammonium assimilation. Metab. Eng. *2*, 69–77.

O'Brien, E.J., and Palsson, B.O. (2015). Computing the functional proteome: recent progress and future prospects for genome-scale models. Curr. Opin. Biotechnol. *34*, 125–134.

O'Brien, E.J., Monk, J.M., and Palsson, B.O. (2015). Using genome-scale models to predict biological capabilities. Cell *161*, 971–987.

Ostergaard, S., Olsson, L., Johnston, M., and Nielsen, J. (2000). Increasing galactose consumption by Saccharomyces cerevisiae through metabolic engineering of the GAL gene regulatory network. Nat. Biotechnol. *18*, 1283–1286.

Österlund, T., Bordel, S., and Nielsen, J. (2015). Controllability analysis of transcriptional regulatory networks reveals circular control patterns among transcription factors. Integr Biol (Camb) *7*, 560–568.

Otero, J.M., Cimini, D., Patil, K.R., Poulsen, S.G., Olsson, L., and Nielsen, J. (2013). Industrial systems biology of Saccharomyces cerevisiae enables novel succinic acid cell factory. PLoS ONE *8*, e54144.

Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M.D., Tai, A., Main, A., Eng, D., et al. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. Nature *496*, 528–532.

Park, J.H., Lee, K.H., Kim, T.Y., and Lee, S.Y. (2007). Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. Proc. Natl. Acad. Sci. USA *104*, 7797–7802.

Park, S.H., Kim, H.U., Kim, T.Y., Park, J.S., Kim, S.-S., and Lee, S.Y. (2014). Metabolic engineering of *Corynebacterium glutamicum* for L-arginine production. Nat. Commun. *5*, 4618.

Patil, K.R., and Nielsen, J. (2005). Uncovering transcriptional regulation of metabolism by using metabolic network topology. Proc. Natl. Acad. Sci. USA *102*, 2685–2689.

Patnaik, R., and Liao, J.C. (1994). Engineering of *Escherichia coli* central metabolism for aromatic metabolite production with near theoretical yield. Appl. Environ. Microbiol. *60*, 3903–3908.

Patnaik, R., Spitzer, R.G., and Liao, J.C. (1995). Pathway engineering for production of aromatics in *Escherichia coli*: Confirmation of stoichiometric analysis by independent modulation of AroG, TktA, and Pps activities. Biotechnol. Bioeng. *46*, 361–370.

Pfleger, B.F., Pitera, D.J., Smolke, C.D., and Keasling, J.D. (2006). Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. Nat. Biotechnol. *24*, 1027–1032.

Qin, J., Zhou, Y.J., Krivoruchko, A., Huang, M., Liu, L., Khoomrung, S., Siewers, V., Jiang, B., and Nielsen, J. (2015). Modular pathway rewiring of Saccharomyces cerevisiae enables high-level production of L-ornithine. Nat. Commun. *6*, 8224.

Redden, H., and Alper, H.S. (2015). The development and characterization of synthetic minimal yeast promoters. Nat. Commun. *6*, 7810.

Redden, H., Morse, N., and Alper, H.S. (2014). The synthetic biology toolbox for tuning gene expression in yeast. FEMS Yeast Res. *15*, 1–10.

Redding-Johanson, A.M., Batth, T.S., Chan, R., Krupa, R., Szmidt, H.L., Adams, P.D., Keasling, J.D., Lee, T.S., Mukhopadhyay, A., and Petzold, C.J. (2011). Targeted proteomics for metabolic pathway optimization: application to terpene production. Metab. Eng. *13*, 194–203.

Regenberg, B., Grotkjaer, T., Winther, O., Fausbøll, A., Åkesson, M., Bro, C., Hansen, L.K., Brunak, S., and Nielsen, J. (2006). Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in *Saccharomyces cerevisiae*. Genome Biol. *7*, R107.

Renata, H., Wang, Z.J., and Arnold, F.H. (2015). Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. Angew. Chem. Int. Ed. Engl. *54*, 3351–3367.

Ro, D.-K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., et al. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature *440*, 940–943.

Rodriguez, A., Kildegaard, K.R., Li, M., Borodina, I., and Nielsen, J. (2015). Establishment of a yeast platform strain for production of p-coumaric acid through metabolic engineering of aromatic amino acid biosynthesis. Metab. Eng. *31*, 181–188.

Saling, P. (2005). Eco-Efficiency Analysis of biotechnological processes. Appl. Microbiol. Biotechnol. *68*, 1–8.

Salis, H.M. (2011). The ribosome binding site calculator. Methods Enzymol. *498*, 19–42.

Salis, H.M., Mirsky, E.A., and Voigt, C.A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. Nat. Biotechnol. *27*, 946–950.

Scalcinati, G., Knuf, C., Partow, S., Chen, Y., Maury, J., Schalk, M., Daviet, L., Nielsen, J., and Siewers, V. (2012). Dynamic control of gene expression in Saccharomyces cerevisiae engineered for the production of plant sesquiterpene α-santalene in a fed-batch mode. Metab. Eng. *14*, 91–103.

Seo, S.W., Kim, D., Latif, H., O'Brien, E.J., Szubin, R., and Palsson, B.O. (2014). Deciphering Fur transcriptional regulatory network highlights its complex role beyond iron metabolism in *Escherichia coli*. Nat. Commun. *5*, 4910.

Seo, S.W., Kim, D., Szubin, R., and Palsson, B.O. (2015). Genome-wide Reconstruction of OxyR and SoxRS Transcriptional Regulatory Networks under Oxidative Stress in *Escherichia coli* K-12 MG1655. Cell Rep. *12*, 1289–1299.

Shi, S., Chen, Y., Siewers, V., and Nielsen, J. (2014). Improving production of malonyl coenzyme A-derived metabolites by abolishing Snf1-dependent regulation of Acc1. MBio *5*, e01130–e14.

Shiba, Y., Paradise, E.M., Kirby, J., Ro, D.K., and Keasling, J.D. (2007). Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. Metab. Eng. *9*, 160–168.

Shih, S.C., Goyal, G., Kim, P.W., Koutsoubelis, N., Keasling, J.D., Adams, P.D., Hillson, N.J., and Singh, A.K. (2015). A versatile microfluidic device for automating synthetic biology. ACS Synth. Biol. *4*, 1151–1164.

Siegel, J.B., Zanghellini, A., Lovick, H.M., Kiss, G., Lambert, A.R., St Clair, J.L., Gallaher, J.L., Hilvert, D., Gelb, M.H., Stoddard, B.L., et al. (2010). Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. Science *329*, 309–313.

Simpson, M.L., Sayler, G.S., Applegate, B.M., Ripp, S., Nivens, D.E., Paulus, M.J., and Jellison, G.E. (1998). Bioluminescent-bioreporter integrated circuits from novel whole-cell biosensors. Trends Biotechnol. *16*, 332–338.

Smolke, C.D., Carrier, T.A., and Keasling, J.D. (2000). Coordinated, differential expression of two genes through directed mRNA cleavage and stabilization by secondary structures. Appl. Environ. Microbiol. *66*, 5399–5405.

Stephanopoulos, G., and Vallino, J.J. (1991). Network rigidity and metabolic engineering in metabolite overproduction. Science *252*, 1675–1681.

Tarca, A.L., Carey, V.J., Chen, X.W., Romero, R., and Drăghici, S. (2007). Machine learning and its applications to biology. PLoS Comput. Biol. *3*, e116. http://dx.doi.org/10.1371/journal.pcbi.0030116.

Thykaer, J., and Nielsen, J. (2003). Metabolic engineering of β-lactam production. Metab. Eng. *5*, 56–69.

Tsao, C.Y., Hooshangi, S., Wu, H.C., Valdes, J.J., and Bentley, W.E. (2010). Autonomous induction of recombinant proteins by minimally rewiring native quorum sensing regulon of E. coli. Metab. Eng. *12*, 291–297.

Tyo, K.E., Ajikumar, P.K., and Stephanopoulos, G. (2009). Stabilized gene duplication enables long-term selection-free heterologous pathway expression. Nat. Biotechnol. *27*, 760–765.

Usaite, R., Jewett, M.C., Oliveira, A.P., Yates, J.R., 3rd, Olsson, L., and Nielsen, J. (2009). Reconstruction of the yeast Snf1 kinase regulatory network reveals its role as a global energy regulator. Mol. Syst. Biol. *5*, 319.

Waks, Z., and Silver, P.A. (2009). Engineering a synthetic dual-organism system for hydrogen production. Appl. Environ. Microbiol. *75*, 1867–1875.

Wang, T., Ma, X., Du, G., and Chen, J. (2012). Overview of regulatory strategies and molecular elements in metabolic engineering of bacteria. Mol. Biotechnol. *52*, 300–308.

Wang, B.L., Ghaderi, A., Zhou, H., Agresti, J., Weitz, D.A., Fink, G.R., and Stephanopoulos, G. (2014). Microfluidic high-throughput culturing of single cells for selection based on extracellular metabolite production or consumption. Nat. Biotechnol. *32*, 473–478.

Westfall, P.J., Pitera, D.J., Lenihan, J.R., Eng, D., Woolard, F.X., Regentin, R., Horning, T., Tsuruta, H., Melis, D.J., Owens, A., et al. (2012). Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. Proc. Natl. Acad. Sci. USA *109*, E111–E118.

Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., Khandurina, J., Trawick, J.D., Osterhout, R.E., Stephen, R., et al. (2011). Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. Nat. Chem. Biol. *7*, 445–452.

Yu, H., and Gerstein, M. (2006). Genomic analysis of the hierarchical structure of regulatory networks. Proc. Natl. Acad. Sci. USA *103*, 14724–14731.

Zhang, F., Carothers, J.M., and Keasling, J.D. (2012). Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. Nat. Biotechnol. *30*, 354–359.

Zhang, Y., Dai, Z., Krivoruchko, A., Chen, Y., Siewers, V., and Nielsen, J. (2015). Functional pyruvate formate lyase pathway expressed with two different electron donors in *Saccharomyces cerevisiae* at aerobic growth. FEMS Yeast Res. *15*, fov024.